ORIGINAL **BELLSOUTH**

EX PARTE OR LATE FILED

Kathleen B. Levitz
Vice President-Federal Regulatory

Suite 900
1133-21st Street, N.W.
Washington, D.C. 20036-3351
202 463-4113
Fax: 202 463-4198
Internet: levitz.kathleen@bsc.bls.com

October 4, 1999

RECEIVED

OCT - 4 1999

FEDERAL COMMUNICATIONS COMMISSION
OFFICE OF THE SECRETARY

Ms. Magalie Roman Salas
Secretary
Federal Communications Commission
The Portals
445 12th St. S.W.
Washington, D.C. 20554

Re: Written Ex Parte in CC Docket No. 98-56 and
CC Docket No. 98-121

Dear Ms. Salas:

This is to inform you that on October 4, 1999 BellSouth Corporation made a written ex parte to Dr. Daniel Shiman of the Common Carrier Bureau's Policy and Program Planning Division. That ex parte consists of a complete copy of Appendix A to a document entitled "Statistical Techniques for the Analysis and Comparison of Performance Measurement Data" that BellSouth filed in the Louisiana Public Service Commission's Docket No. U-22252-Subdocket C. This filing was the subject of a written ex parte filing BellSouth made on September 30, 1999, in which only the first five pages of Appendix A were included. Today's filing assures that the copy of Appendix A in the record is complete. Copies of Appendix A have also been sent to Florence Setzer, Alex Belinfante, Andre Rausch, and Whitey Thayer, all members of the Common Carrier Bureau staff.

Pursuant to Section 1.1206(a)(1) of the Commission's rules, I am filing two copies of this notice and that written ex parte presentation in both the dockets identified above. Please associate this notification with the record in both those proceedings.

Sincerely,

Kathleen B. Levitz

Kathleen B. Levitz

Attachment

cc: Daniel Shiman (w/o attachment)
Florence Setzer (w/o attachment)
Alex Belinfante (w/o attachment)
Andre Rausch (w/o attachment)
Whitey Thayer (w/o attachment)

BELLSOUTH

Kathleen B. Levitz
Vice President-Federal Regulatory

Suite 900
1133-21st Street, N.W.
Washington, D.C. 20036-3351
202 463-4113
Fax: 202 463-4198
Internet: levitz.kathleen@bsc.bls.com

October 4, 1999

Dr. Daniel Shiman
Policy and Program Planning Division
Common Carrier Bureau
Federal Communications Commission
1919 M Street, N.W.
Washington, D.C. 20554

Written Ex Parte in CC Docket No. 98-121 and CC Docket No. 98-56

Dear Dr. Shiman:

On September 30, 1999, I sent you two documents that BellSouth filed recently in the Louisiana Public Service Commission's proceeding LPSC Docket Number U-22252-C. Inadvertently the last four pages of Appendix A to the document entitled "Statistical Techniques for the Analysis and Comparison of Performance Measurement Data" were not included in that earlier submission. Attached is a copy of Appendix A that includes the missing pages. If after reviewing this attachment you conclude that you need additional information, please call me at 202.463.4113.

In compliance with Section 1.1206(b)(1) of the Commission's rules, I have filed with the Secretary of the Commission two copies of this written ex parte presentation for inclusion in the records of both CC Docket No. 98-56 and CC Docket No. 98-121.

Sincerely,

Kathleen B. Levitz

Kathleen B. Levitz

Attachment

cc: Alex Belinfante
Florence Setzer
Whitey Thayer
Andre Rausch

Appendix A. The Truncated Z Statistic

The Truncated Z test statistic was developed by Dr. Mallows in order to have an aggregate level test when transaction level data are available that

- provides a single overall index on a standard scale;
- is robust with respect to unnecessary disaggregation,
- incorporates the number of observations in a cell into the determination of the weight for the contribution of each comparison cell,
- limits the amount of "neutralization" between comparison cells, and
- is a continuous function of the observations.

The Ernst & Young statistical team and Dr. Mallows have studied the implementation of the statistic using some of BellSouth's performance measure data. This has resulted in an overall process for comparing CLEC an ILEC performance such that the following principles hold:

1) Like-to-Like Comparisons are made. (See Appendix B for an example based on the trunk blocking measure.)
2) Error probabilities are balanced. (See Appendix C)
3) Extreme values are trimmed from the data sets when they significantly distort the performance measure statistic. (See Appendix E)
4) The testing process is an automated production system. (Discussed here. See Appendix D for reporting guidelines.)
5) The determination of ILEC favoritism is based on a single aggregate level test statistic. (Discussed here.)

This appendix provides the details behind computing the Truncated Z test statistic so that principles 4 and 5 hold. We start by assuming that any necessary trimming of the data is complete, and that the data are disaggregated so that comparisons are made within appropriate classes or adjustment cells that define "like" observations.

Notation and Exact Testing Distributions

Below, we have detailed the basic notation for the construction of the truncated z statistic. In what follows the word "cell" should be taken to mean a like-to-like comparison cell that has both one (or more) ILEC observation and one (or more) CLEC observation.

L = the total number of occupied cells

j = $1,\ldots,L$; an index for the cells

n_{1j} = the number of ILEC transactions in cell j

n_{2j} = the number of CLEC transactions in cell j

n_j = the total number transactions in cell j; $n_{1j} + n_{2j}$

X_{1jk} = individual ILEC transactions in cell j; $k = 1,\ldots, n_{1j}$

X_{2jk} = individual CLEC transactions in cell j; $k = 1,\ldots, n_{2j}$

Y_{jk} = individual transaction (both ILEC and CLEC) in cell j

$$= \begin{cases} X_{1jk} & k = 1,\ldots, n_{1j} \\ X_{2jk} & k = n_{1j}+1,\ldots, n_j \end{cases}$$

$\Phi^{-1}(\cdot)$ = the inverse of the cumulative standard normal distribution function

For Mean Performance Measures the following additional notation is needed.

$\overline{X}_{1j}$ = the ILEC sample mean of cell j

$\overline{X}_{2j}$ = the CLEC sample mean of cell j

s^2_{1j} = the ILEC sample variance in cell j

s^2_{2j} = the CLEC sample variance in cell j

y_{jk} = a random sample of size n_{2j} from the set of $Y_{j1},\ldots, Y_{jn_j}$; $k = 1,\ldots,n_{2j}$

M_j = the total number of distinct pairs of samples of size n_{1j} and n_{2j};

$$= \binom{n_j}{n_{1j}}$$

The exact parity test is the permutation test based on the "modified Z" statistic. For large samples, we can avoid permutation calculations since this statistic will be normal (or Student's t) to a good approximation. For small samples, where we cannot avoid permutation calculations, we have found that the difference between "modified Z" and the textbook "pooled Z" is negligible. We therefore propose to use the permutation test based on pooled Z for small samples. This decision speeds up the permutation computations considerably, because for each permutation we need only compute the sum of the CLEC sample values, and not the pooled statistic itself.

A permutation probability mass function distribution for cell j, based on the "pooled Z" can be written as

$$PM(t) = P(\sum_k y_{jk} = t) = \frac{\textit{the number of samples that sum to } t}{M_j},$$

and the corresponding cumulative permutation distribution is

$$CPM(t) = P(\sum_k y_{jk} \le t) = \frac{\textit{the number of samples with sum} \le t}{M_j}.$$

For Proportion Performance Measures the following notation is defined

a_{1j} = the number of ILEC cases possessing an attribute of interest in cell j

a_{2j} = the number of CLEC cases possessing an attribute of interest in cell j

a_j = the number of cases possessing an attribute of interest in cell j; $a_{1j}+ a_{2j}$

The exact distribution for a parity test is the hypergeometric distribution. The hypergeometric probability mass function distribution for cell j is

$$HG(h) = P(H = h) = \begin{cases} \dfrac{\binom{n_{1j}}{h}\binom{n_{2j}}{a_j - h}}{\binom{n_j}{a_j}}, & \max(0, a_j - n_{2j}) \le h \le \min(a_j, n_{1j}) \\ 0 & \text{otherwise} \end{cases},$$

and the cumulative hypergeometric distribution is

$$CHG(x) = P(H \le x) = \begin{cases} 0 & x < \max(0, a_j - n_{1j}) \\ \sum_{h=\max(0,a_j-n_{1j})}^{x} HG(h), & \max(0, a_j - n_{1j}) \le x \le \min(a_j, n_{2j}) \\ 1 & x > \min(a_j, n_{2j}) \end{cases}.$$

For Rate Measures, the notation needed is defined as

b_{1j} = the number of ILEC base elements in cell j

b_{2j} = the number of CLEC base elements in cell j

b_j = the total number of base elements in cell j; $b_{1j}+ b_{2j}$

$\hat{r}_{1j}$ = the ILEC sample rate of cell j; n_{1j}/b_{1j}

$\hat{r}_{2j}$ = the CLEC sample rate of cell j; n_{2j}/b_{2j}

q_j = the relative proportion of CLEC elements for cell j; b_{2j}/b_j

The exact distribution for a parity test is the binomial distribution. The binomial probability mass function distribution for cell j is

$$BN(k) = P(B = k) = \begin{cases} \binom{n_j}{k} q_j^k (1-q_j)^{n_j-k}, & 0 \le k \le n_j \\ 0 & \text{otherwise} \end{cases},$$

and the cumulative binomial distribution is

$$CBN(x) = P(B \le x) = \begin{cases} 0 & x < 0 \\ \sum_{k=0}^{x} BN(k), & 0 \le x \le n_j \\ 1 & x > n_j \end{cases}.$$

Calculating the Truncated Z

The general methodology for calculating an aggregate level test statistic is outlined below.

1. **Calculate cell weights,** W_j. A weight based on the number of transactions is used so that a cell which has a larger number of transactions has a larger weight. The actual weight formulae will depend on the type of measure.

Mean Measure

$$W_j = \sqrt{\frac{n_{1j} n_{2j}}{n_j}}$$

Proportion Measure

$$W_j = \sqrt{\frac{n_{2j} n_{1j}}{n_j} \cdot \frac{a_j}{n_j} \cdot \left(1 - \frac{a_j}{n_j}\right)}$$

Rate Measure

$$W_j = \sqrt{\frac{b_{1j} b_{2j}}{b_1} \cdot \frac{n_j}{b_j}}$$

2. **In each cell, calculate a Z value,** Z_j. A standard normal Z statistic is needed for each cell.

- If $W_j = 0$, set $Z_j = 0$.
- When the cell sample sizes are sufficiently large, formulae based on a normal approximation can be used.
- If cell sample sizes are not large enough for a normal approximation to hold, then exact testing methods must be employed. When this occurs, the results of the test statistic are converted into an equivalent value from the standard normal distribution.

The actual Z statistic calculation depends on the type of performance measure.

Mean Measure

$$Z_j = \Phi^{-1}(\alpha)$$

where α is determine by the following algorithm.

If $\min(n_{1j}, n_{2j}) > 6$, then determine α as

$$\alpha = P(t_{n_{1j}-1} \leq T_j),$$

that is, α is the probability that a t random variable,

$$t_j = \frac{\bar{X}_{1j} - \bar{X}_{2j}}{s_{1j}\sqrt{\frac{1}{n_{1j}} + \frac{1}{n_{2j}}}}$$

with n_{1j} - 1 degrees of freedom, is less than T_j.

$$T_j = t_j + \frac{g}{6}\left(\frac{n_{1j} + 2n_{2j}}{\sqrt{n_{1j}\, n_{2j}(n_{1j} + n_{2j})}}\right)\left(t^2 + \frac{n_{2j} - n_{1j}}{2n_{1j} + n_{2j}}\right).$$

Here the coefficient g is an estimate of the skewness of the parent population, which we assume is the same in all cells. It can be estimated from the ILEC values in the largest cells. This needs to be done only once for each measure. We have found that attempting to estimate this skewness parameter for each cell separately leads to excessive variability in the "adjusted" t. We therefore use a single compromise value in all cells.

Note, that t_j is the "modified Z" statistic. The statistic T_j is a "modified Z" corrected for the skewness of the ILEC data.

If $\min(n_{1j}, n_{2j}) \le 6$, and

a) $M_j \le 1{,}000$ (the total number of distinct pairs of samples of size n_{1j} and n_{2j} is 1,000 or less).

- Calculate the sample sum for all possible samples of size n_{2j}.
- Rank the sample sums from smallest to largest. Ties are dealt by using average ranks.
- Let R_0 be the rank of the observed sample sum with respect all the sample sums.

$$\alpha = 1 - \frac{R_0 - 0.5}{M_j}$$

b) $M_j > 1{,}000$

- Draw a random sample of 1,000 sample sums from the permutation distribution.
- Add the observed sample sum to the list. There is a total of 1001 sample sums. Rank the sample sums from smallest to largest. Ties are dealt by using average ranks.
- Let R_0 be the rank of the observed sample sum with respect all the sample sums.

$$\alpha = 1 - \frac{R_0 - 0.5}{1001}.$$

Proportion Measure

$$Z_j = \frac{n_j\, a_{1j} - n_{1j}\, a_j}{\sqrt{\dfrac{n_{1j}\, n_{2j}\, a_j\,(n_j - a_j)}{n_j - 1}}}.$$

Rate Measure

If $n_{1j}>15$, $n_{2j}>15$, and $n_j q_j (1 - q_j) > 9$ then

$$Z_j = \frac{\hat{r}_{1j} - \hat{r}_{2j}}{\sqrt{\hat{r}_{1j}\left(\dfrac{1}{b_{1j}} + \dfrac{1}{b_{2j}}\right)}}.$$

Otherwise,

$$Z_j = \Phi^{-1}(\alpha)$$

where

$$\alpha = 1 - CBN(n_{2j}-1).$$

3. **Obtain a truncated Z value for each cell,** Z_j^*. To limit the amount of cancellation that takes place between cell results during aggregation, cells whose results suggest possible favoritism are left alone. Otherwise the cell statistic is set to zero. This means that positive equivalent Z values are set to 0, and negative values are left alone. Mathematically, this is written as

$$Z_j^* = \min(0, Z_j).$$

4. **Calculate the theoretical mean and variance of the truncated statistic under the null hypothesis of parity,** $E(Z_j^*|H_0)$ and $Var(Z_j^*|H_0)$. In order to compensate for the truncation in step 3, an aggregated, weighted sum of the Z_j^* will need to be centered and scaled properly so that the final aggregate statistic follows a standard normal distribution.

   - If $W_j = 0$, then no evidence of favoritism is contained in the cell. The formulae for calculating $E(Z_j^* \mid H_0)$ and $Var(Z_j^* \mid H_0)$ cannot be used. Set both equal to 0.

   - If the equivalent Z value of a mean or rate measure was calculated using a normal approximation, or $\min(n_{1j}, n_{2j}) > 30$ for a proportion measure then

$$E(Z_j^* \mid H_0) = -\frac{1}{\sqrt{2\pi}}, \text{ and}$$

$$Var(Z_j^* \mid H_0) = \frac{1}{2} - \frac{1}{2\pi}.$$

   - Otherwise, determine the total number of values for Z_j^*, denoted by N_j. Let z_{ji} and θ_{ji}, $i = 1,\ldots,N_j$, denote the values of Z_j^* and the probabilities of observing each value, respectively.

$$E(Z_j^* \mid H_0) = \sum_{i=1}^{N_j} \theta_{ji} z_{ji}, \text{and}$$

$$\mathrm{Var}(Z_j^* \mid H_0) = \sum_{i=1}^{N_j} \theta_{ji} z_{ji}^2 - \left[E(Z_j^* \mid H_0)\right]^2.$$

The actual values of the z's and θ's depends on the type of measure.

Mean Measure

$$N_j = \min(M_j, 1{,}000)$$

$$z_{ji} = \min\left\{0, 1 - \Phi^{-1}\left(\tfrac{R_i - 0.5}{N_j}\right)\right\} \quad \text{where } R_i \text{ is the rank of sample sum i}$$

$$\theta_j = \frac{1}{N_j}$$

Proportion Measure

$$N_j = \min(a_j, n_{2j}) - \max(0, a_j - n_{1j}) + 1$$

$$z_{ji} = \min\left\{0, \frac{n_j\, i - n_{1j}\, a_j}{\sqrt{\dfrac{n_{1j}\, n_{2j}\, a_j\, (n_j - a_j)}{n_j - 1}}}\right\}, \quad i = 0, K, a_j$$

$$\theta_{ji} = HG(i)$$

Rate Measure

$$N_j = n_j$$

$$z_{ji} = \min\{0, \Phi^{-1}(1 - CBN(i-1))\}$$

$$\theta_{ji} = BN(i)$$

5. **Calculate the aggregate test statistic, Z^T.**

$$Z^T = \frac{\sum_j W_j Z_j^* - \sum_j W_j E(Z_j^* \mid H_0)}{\sqrt{\sum_j W_j^2 \mathrm{Var}(Z_j^* \mid H_0)}}$$

Decision Process

Once Z^T has been calculated, it is compared to a critical value to determine if the ILEC is favoring its own customers over a CLEC's customers. The derivation of the critical value is found in Appendix C.

This critical value changes as the ILEC and CLEC transaction volume change. One way to make this transparent to the decision maker, is to report the difference between the test statistic and the critical value, *diff* = $Z^T - c_B$. If favoritism is concluded when $Z^T < c_B$, then the *diff* < 0 indicates favoritism.

This make it very easy to determine favoritism: a positive *diff* suggests no favoritism, and a negative *diff* suggests favoritism. Appendix D provides an example of how this information can be reported for each month.